Floor Discussion

Richard Moore
Bureau of the Census

The first paper, presented by Leon Willenbourg of Statistics Netherlands, discussed recent developments for the ARGUS interactive disclosure limitation software for microdata. A successor version (μ-ARGUS) now has automated editing capabilities (no interaction is required). In addition, Statistics Netherlands is developing software for tabular products (τ-ARGUS).

Timothy Evans of the Bureau of the Census gave the second presentation. It discussed some preliminary experimental results for the use of multiplicative noise to mask economic tabulations. Use of such a technique may require certain supplemental procedures such as some complimentary suppression and/or iterative proportional fitting.

Stephen Fienberg of Carnegie Mellon University presented the final paper. He related some of his ideas for a novel approach for the construction of a public use microdata file. This approach uses a log-linear model to generate artificial data based on characteristics of the original , "cleaned up" information.

Carolyn Shettle began the floor questioning. She asked Leon Willenbourg about his criterions for risk in microdata. He replied that an individual was considered at risk of re-identification if less than 100 people in the entire population had a similar combination of traits.

Leon Willenbourg was also asked about additional risk associated with an automated masking process. Could an algorithm be developed to unmask the file? Mr. Willenbourg replied that Statistics Netherlands is aware of such problems, especially if variables are inter-dependent.

Fritz Scheuren asked Tim Evans about the effect of the multiplicative noise technique on time-series analysis. Mr. Evans replied that each multiplier would be changed only slightly from period to period.

Mr. Scheuren also asked Tim Evans whether he intended to rake (iteratively proportional fit) each table separately. Mr. Evans replied that research was focusing on a procedure that coordinated the raking procedure into one operation.

Mr. Scheuren asked Stephen Fienberg about the reliability of the resampling process. Mr. Fienberg stated that the process was sound enough to allow the user to go from the data to information to wisdom.

Myron Straf inquired about whether Stephen Fienberg's procedure was dependent on the agency's perception of response versus independent variables. Mr. Fienberg felt that it is necessary for agencies to make much, if not all, the data available to the users.

Brian Greenberg asked the panelists about the relationship between disclosure and anticipated needs. Mr. Fienberg replied that agencies have a good feel for current user's needs. He wondered whether agencies will be able to predict the future needs.